基于自适应门控图神经网络的交通流预测 *

王 杨，郑 津†，刘 影，李 平

(西南石油大学 计算机科学学院，成都 610500)

摘 要：交通流预测是智能交通系统中的重要组成部分，由于交通数据的复杂性，长期而又准确的交通流预测一直是时间序列预测中最具挑战性的任务之一。近年来，研究人员将基于图神经网络的时空图建模方法应用于交通流预测任务，并取得了良好的预测性能。然而，现有的图建模方法仅通过预定义的邻接结构反映道路网络中的空间依赖关系，忽略了各节点之间的序列关联关系对预测的重要性。针对这一局限性，提出了一种自适应门控图神经网络(Ada-GGNN)，其核心为通过空间传递模块同时捕获道路网络的空间结构及自适应的时序相关性，并通过门控机制学习节点上的时间序列特征。在两个真实交通网络数据集 PeMSD7 和 Los-loop 上的实验结果证明了该模型具有更优越的性能。

关键词：交通流预测；时空图；自适应门控图神经网络；时序相关性

中图分类号：TP391 doi: 10.19734/j.issn.1001-3695.2022.01.0017

Traffic flow prediction based on adaptive gated graph neural network

Wang Yang, Zheng Jin†, Liu Ying, Li Ping

(*College of Computer Science*, *Southwest Petroleum University*, *Chengdu* 610500, *China*)

Abstract: Traffic flow prediction is an important part of intelligent transportation system. Due to the complexity of traffic data, long-term and accurate traffic flow prediction has always been one of the most challenging tasks in time series forecasting. In recent years, researchers have applied spatial-temporal graph modeling methods based on graph neural networks to traffic flow prediction tasks which achieved good prediction performance. However, existing graph modeling methods only reflect the spatial dependence in road networks through predefined adjacency structures, ignoring the importance of time-series correlation between nodes for prediction. Aiming at this limitation, this paper proposed an Adaptive Gated Graph Neural Network (Ada-GGNN) , the core of which was to simultaneously capture the spatial structure of the road network and adaptive time series correlation through the spatial passing module, and learned time-series features on nodes through the gating mechanism. The experimental results on two real-world traffic network datasets PeMSD7 and Los-loop show that the model has better performance.

Key words: traffic flow prediction; spatial-temporal graph; adaptive gated graph neural networks; time series correlation

0 引言

智能交通系统是智慧城市发展不可或缺的组成部分，而准确实时的交通流预测作为智能交通系统的基础和重要研究方向，它是减少交通事故、缓解交通拥堵和提高交通运输效率的有效解决方案。在实际生活中，交通流通常由道路上的传感器记录，如流量、车速、道路占用率等。交通流预测旨在根据道路网络中的历史交通数据(如传感器记录)预测未来一段时间的交通数据。传统的交通流预测模型主要是基于统计或机器学习的方法，如历史平均模型(HA)[1]、自回归整合移动平均模型(ARIMA)[2]、向量自回归(VAR)[3]、k-最近邻模型(KNN)[4]等，以上方法在实际应用中往往预测效果不佳，它们依赖于数据的平稳性假设。然而交通状况是复杂的，受各种环境因素影响，表现出随机性、周期性、趋势性和时空性等特点。其次，这些方法需要经过复杂的特征工程进行建模，并且只考虑了序列的时间信息，而忽略了空间信息对于交通流预测的重要性。

近年来，深度学习因其强大的特征表达能力和非线性拟合能力被广泛应用于交通流预测任务。Yu 等人[5]使用深度的长短期记忆网络(LSTM)来预测交通序列，实验证明了基于神经网络的方法相较于传统方法具有更优异的结果，但是该方法仍没有考虑交通路网的空间结构。Shi 等人[6]设计了 ConvLSTM 模型用于降雨预测，它将卷积神经网络(CNN)和循环神经网络(RNN)进行结合，分别对空间和时间相关性进行建模。之后，Zhang 等人[7]将交通路网转换为规则的网格，然后采用 CNN 提取空间特征。虽然这些模型可以提取数据的时空特征，但交通道路本质上是不规则的，将交通数据视为规则的 2 维或 3 维网格数据会导致空间拓扑信息的丢失。

随后，研究人员引入了图神经网络(GNN)来对时空网络数据进行建模，并在交通流预测任务上取得了巨大进步。Li 等人[8]提出了 DCRNN 模型，它将交通网络建模为有向图，采用双向随机游走捕获道路之间的空间依赖性，并使用 RNN 捕获时间相关性，该方法同时处理了时间和空间信息，具有更准确的预测结果。此后，研究人员对 GNN 与 CNN 或 RNN 结合起来的这种研究范式进行了深入挖掘，例如：文献[9~11]都采用了 GNN 来获取拓扑结构中的空间依赖性，采用一维卷积或 RNN 来获取动态的时间依赖性。

尽管这些基于图神经网络的方法在交通流预测任务上已取得相对成功，但这些模型仍然存在缺陷。首先，这些方法仅采用预定义的图邻接结构捕获节点间的空间依赖关系，却

收稿日期：2022-01-18；**修回日期：**2022-03-10 **基金项目：**国家杰出青年科学基金资助项目(61625204)；四川省科技计划资助项目(2020YJ0125)；西南石油大学科研创新能力提升计划“启航”项目(2019QHZ016)

作者简介：王杨(1995-)，男，重庆綦江人，硕士研究生，主要研究方向为图神经网络；郑津(1984-)，男(通信作者)，四川广安人，讲师，博士研究生，主要研究方向为深度学习、图神经网络(zhengjin@swpu.edu.cn)；刘影(1980-)，女，山东烟台人，副教授，硕导，博士，主要研究方向为复杂网络、网络传播动力学、数据挖掘；李平(1977-)，女，四川成都人，教授，硕导，博士，主要研究方向为图神经网络、自然语言处理、复杂网络.

忽略了各节点上时序之间的相关性。预定义的图邻接结构通常基于传感器地理位置之间的距离或上下游关系构建邻接矩阵，但这样会导致两个节点上时间序列相关却无直接连边的情况发生。如图 1 所示，左图表示交通路网中的节点分布，右图中实线表示预定义的邻接结构，虚线表示存在时序相关性的节点对。从图中可以看出，仅利用预定义的邻接关系无法全面反映节点之间的相互影响。其次，现有图神经网络建模交通流预测方法的另一个缺陷在于，这类方法往往分别探索时间和空间维度的特征，最后再进行融合，这种延迟融合策略会导致模型不能实时获取全局信息，从而影响模型预测效果。

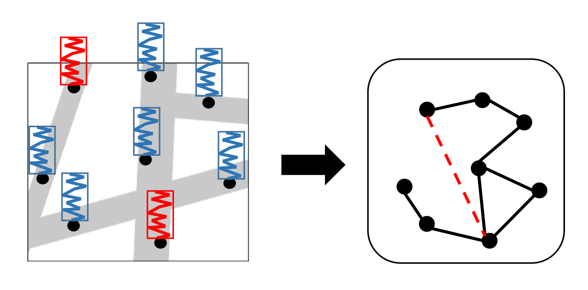


图 1　问题描述

Fig. 1　Problem description

为解决上述问题，本文提出了一种新的端到端的时空数据预测模型，即 Ada-GGNN。首先，该模型通过空间传递模块捕获交通路网的空间依赖性，并从中挖掘出节点间时序的关联关系。然后通过 GRU[12]学习交通数据动态的时间依赖性，并在每个时间步进行多次的时空融合，探索时间和空间的高阶交互，获得时空嵌入表示，最后将这些表示通过全连接层进行最后的嵌入，得到预测结果。该工作的主要贡献如下：

(1)通过一种可学习的方式来获取时序间的自适应相关性矩阵，以捕获节点间因预定义邻接结构而丢失的关联关系。

(2)采用循环聚合方法来实时融合时空嵌入表示，使得模型的时空嵌入表示更丰富。

(3)在两个真实交通数据集上，Ada-GGNN 的实验结果证明了该方法的有效性。同时，所提出的方法具有很好的泛化性，很容易拓展到其他时空数据挖掘任务中。

1　相关工作

1.1　图神经网络

Gori 等人[13]首次提出了图神经网络(GNN)，并将其应用于图结构数据。随后，关于 GNN 的改进及其变体在各项任务中取得了极大的进步。Defferrard 等人[14]采用一种快速局部卷积滤波器对 GCN 进行了改进，提出了 ChebNet 模型。Kipf 等人[15]使用基于光谱卷积的一阶近似来简化 ChebNet 模型，并在半监督分类任务中达到了极好的分类效果。文献[16]通过采样每个节点固定数量的邻居并进行聚合，从而提高了大图的可扩展性。文献[17]提出了 GAT 模型，它是一种强大的 GCN 变体，通过引入一种注意力机制来动态地确定每个邻居节点对中心节点的重要性。然而，以上的 GNN 模型通常采用多层感知机更新节点的隐藏状态，因此，在图的长范围信息传播中存在限制。

为了解决上述问题，Li 等人[18]提出了门控图神经网络(GGNN)，这是一种基于 GRU 的经典空间域消息传递模型。由于 GGNN 在时空数据上的有效性，它已被应用于各种任务。Zhang 等人[19]提出了 TextING 模型，该模型为每个文档构建单独的图，然后使用 GGNN 来学习词节点的嵌入，通过大量实验，该方法已被证明优于最先进的文本分类方法。文献[20]专注于一种新颖的金融事件预测任务，它使用 GGNN 基于事件图来更新事件表示。文献[21]改进了 GGNN 模型以更好地模拟节点交互并更好地从图中推断出兼容性。此外，GGNN 还被应用于其他任务，如推荐[22]、图像分类[23]、情景识别[24]等。对于时空数据预测问题，GGNN 表现出了优秀的性能，受此启发，本文采用 GGNN 框架和可学习的自适应节点相关性矩阵来共同应对长短期的交通预测问题。

2　模型

2.1　问题定义

在交通预测任务中通常将每个传感器看做一个节点，传感器之间的距离或上下游关系构成节点之间的连边，则交通网络可被定义为无向图 $G=(V,E,\mathbb{A})$，其中 V 是节点集合，E 是节点间边的集合，$\mathrm{A}\in\mathbb{R}^{N\times N}$ 是图 G 的邻接矩阵。在时间步 t 时，N 个节点的交通状况表示为 $\mathrm{X}^{t}\in\mathbb{R}^{N\times D}$，其中 D 表示特征维度。因此，交通预测问题可以看做是给定一个路网拓扑结构 G 和 T 个时间步的历史交通信息，学习一个映射函数 f，使得能够预测接下来 τ 个时刻的交通信息，其定义如式(1)所示。

$$[\mathrm{X}^{t+1},\cdots,\mathrm{X}^{t+\tau}]=f([\mathrm{X}^{t-T+1},\cdots,\mathrm{X}^{t-1},\mathrm{X}^{t}],G) \tag{1}$$

2.2　模型总体架构

Ada-GGNN 模型主要由每个时间步的两阶段时空融合模块(ST Fusion module)组成，每个时空融合模块包含两个部分：空间传递模块(SP module，Spatial Passing module)和门控循环单元(GRU)。如图 2 所示，首先使用 T 个时间步的历史数据作为输入，在每个时间步中，利用空间传递模块捕获交通路网的固定拓扑结构和学习序列之间的自适应关联关系。然后，将具有空间特征的时间序列输入到门控循环单元，通过门控循环单元之间的信息传递获取序列数据的动态变化，捕捉时间特征并进行时空数据融合。最后，通过一个全连接层得到 τ 个时间步的预测结果。

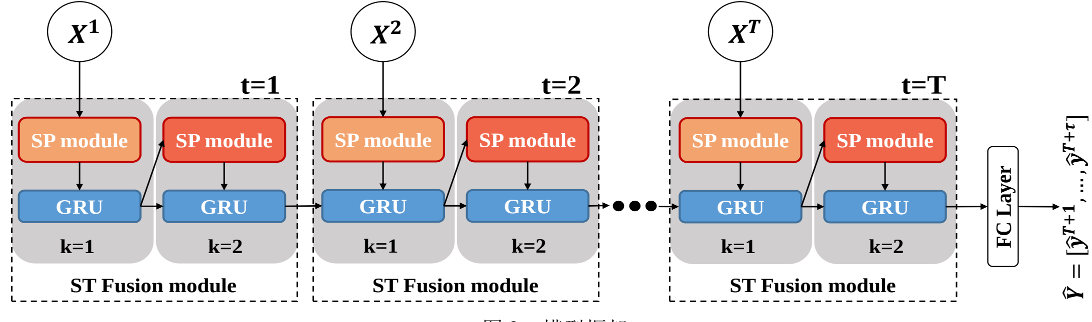


图 2　模型框架

Fig. 2　Model architecture diagram

2.3　空间传递模块

考虑到交通路网的拓扑结构，空间传递模块主要用来捕获道路之间的空间依赖性关系。如图 3 所示，空间传递模块包含两部分，一方面是通过预定义的邻接矩阵获取固定的空间关系，另一方面通过可学习的相关性矩阵获取节点间的关联关系。在该模块中，采用图卷积操作对节点间传递的信息进行提取。对于交通路网中显式的空间关系，通常基于传感器之间的位置距离构建其邻接矩阵，表示为 $\mathrm{A}_{\mathrm{org}}\in\mathbb{R}^{N\times N}$。令

$I^{t,k} \in \mathbb{R}^{N\times F}$ 表示输入的交通数据，$W_{org}^{k} \in \mathbb{R}^{F\times D}$ 表示模型参数矩阵，并采用 ReLU 激活函数，因此该图卷积层定义如式(2)所示。

$$Z_{org}^{t,k} = ReLU\left(A_{org} I^{t,k} W_{org}^{k}\right), k \in [1,2] \quad (2)$$

对于交通路网中节点上时间序列的关联关系挖掘，令 $A_{ada} \in \mathbb{R}^{N\times N}$ 表示可学习的节点间自适应相关性矩阵，并随机初始化，通过模型的训练自动捕获和调整节点间关联关系的大小。该图卷积层定义如式(3)所示。

$$Z_{ada}^{t,k} = ReLU\left(A_{ada} I^{t,k} W_{ada}^{k}\right), k \in [1,2] \quad (3)$$

值得注意的是，该模块对交通路网中的拓扑关系进行了二阶交互($k \in [1,2]$)，因此，当 k 为 1 时，$I^{t,k=1} = X^{t}$ ；k 为 2 时，$I^{t,k=2} = GRU^{t,k=1}$ (如图 1 所示)。最后，将两部分提取到的信息通过拼接的方式进行特征融合，如式(4)所示。

$$Z^{t,k} = Z_{org}^{t,k} \,||\, Z_{ada}^{t,k}, k \in [1,2] \quad (4)$$

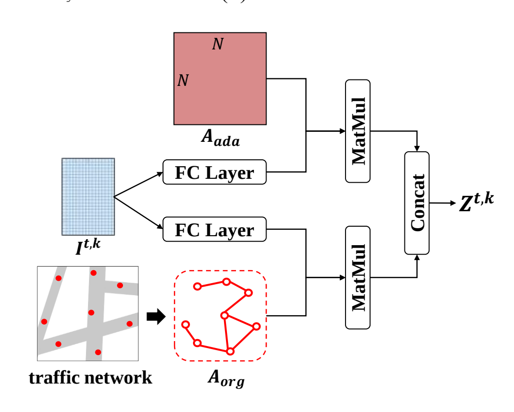


图 3　空间传递模块结构

Fig. 3　The architecture of the SP module

2.4　时空融合模块

空间传递模块捕获了每个时间步节点之间的空间特征和序列相关性，并提取了它们融合后的特征表达，但交通预测是一个时空预测任务，不仅仅要捕获空间信息，还需要捕获序列上的时间依赖性。因此，在时空融合模块，通过 GRU 捕获序列时间维度上的依赖性，并以循环聚合的方式在每个时间步进行实时的时空融合。如公式(5)所示，GRU 接收两个输入，一个是具有空间信息的当前时刻数据嵌入表示 $Z^{t,k}$，另一个是上一状态表示 H_{pre}。由于在每个时间步 Ada-GGNN 采取了二阶的时空交互(如图 1 所示)，所以在 k 阶时的上一状态表达有所不同，具体如公式(6)所示。

$$H^{t,k} = GRU\left(Z^{t,k}, H_{pre}\right) \quad (5)$$

$$\begin{cases} H_{pre} = H^{t-1,k+1}, k=1 \\ H_{pre} = H^{t,k-1}, k=2 \end{cases} \quad (6)$$

GRU 通过门控机制来控制当前时刻输入和上一状态的写入程度，更新门的值越大，重置门的值越小，则表示上一状态的信息被写入的越少。以这种方式，Ada-GGNN 能够捕获交通数据长期的时间相关性，其详细操作如下公式所示。

$$u^{t,k} = \sigma\left(W_u\left[Z^{t,k}, H_{pre}\right] + b_u\right) \quad (7)$$

$$r^{t,k} = \sigma\left(W_r\left[Z^{t,k}, H_{pre}\right] + b_r\right) \quad (8)$$

$$\tilde{h}^{t,k} = tanh\left(W_{\tilde{h}}\left[Z^{t,k}, r^{t,k} * H_{pre}\right] + b_{\tilde{h}}\right) \quad (9)$$

$$H^{t,k} = \left(1 - u^{t,k}\right) * H_{pre} + u^{t,k} * \tilde{h}^{t,k} \quad (10)$$

其中，$u^{t,k}$ 和 $r^{t,k}$ 分别表示 GRU 中更新门和重置门大小，σ 表示 Sigmoid 激活函数，W 和 b 是可学习参数。

2.5　预测

经过 T 个时间步的时空融合后，Ada-GGNN 采用全连接层得到最后的预测结果 $\hat{Y}$，如公式(12)所示。

$$\hat{Y} = W_{out} H^{t=T,k=2} + b_{out} \quad (11)$$

其中，W_{out} 和 b_{out} 为全连接层参数。考虑到交通预测问题是一个回归任务，所以采用平均绝对误差(MAE)作为损失函数，其定义如公式(13)所示。

3　实验

3.1　数据集

为了评估所提出模型的有效性，在两个公开可用的真实交通数据集 PeMSD7 和 Los-loop 上进行了实验。PeMSD7 数据集是加利福尼亚州 2012 年 5 月和 6 月(仅工作日，共 44 天)的交通数据，由 228 个传感器收集的交通速度组成。Los-loop 数据集是洛杉矶高速公路上 207 个检测器收集的 2012 年 3 月 1 日至 7 日的交通速度数据。在所有实验中，将每间隔 5 分钟汇总一次的交通速度作为交通特征，并选择前 80%的数据作为训练集，剩下的 20%作为测试集。其数据集的具体统计信息如表 1 所示。

表 1　两个数据集的统计信息

Tab. 1　Summary statistics of two traffic network datasets

数据集	节点	边	时间步长	最大值	平均值
PeMSD7	228	832	12672	82.6	58.9
Los-loop	207	2833	2016	70.0	58.9

3.2　实验设置

在实验中，采用 Python 编程语言和 Pytorch 深度学习框架对模型进行实现，将前 60 分钟的交通数据作为模型的输入预测未来 15、30、60 分钟的交通速度。训练过程中，学习率设置为 0.001，epoch 设置为 1000，采用 Adam 优化器[25]。为防止模型过拟合，采用“早停法”(Early Stopping)进行训练。所有实验均在具有 16GB RAM 的单个 NVIDIA Tesla T4 GPU 上进行。

本文所提出的模型与以下七种基线方法进行比较：

a) ARIMA：自回归整合移动平均模型是一种经典的自回归模型，它将自回归、滑动平均以及差分法结合在一起对时间序列数据进行预测。

b) SVR[26]：支持向量回归模型是 SVM 的重要应用分支，实验中采用线性核函数进行交通预测任务。

c) ASTGCN[10]：该模型主要由三个独立的组件组成，分别对小时、天、周的时间属性进行建模，每个组件都包含了时空注意力机制和时空卷积两个操作部分，并通过加权融合的方式产生最终的预测。

d) T-GCN[11]：该模型通过 GCN 学习交通路网的拓扑结构，GRU 来学习交通数据的动态变化，从而捕获空间依赖性和时间依赖性。

e) A3T-GCN[27]：A3T-GCN 在 T-GCN 的基础上，引入了注意力机制来调整不同时间点的重要性并组装全局时间信息来进行交通预测。

f) LSGCN[28]：该模型将 GCN 和图注意力网络集成到一个空间门控块中，通过空间门控块和门控线性单元[29](GLU)来捕获数据的时空特征并得到预测结果。

g) GWN[30]：该模型通过一个可学习的方式来捕捉数据中隐藏的空间依赖，同时，使用扩张一维卷积组件扩大感受野以此来处理长序列数据。

为了验证 Ada-GGNN 的有效性，采用以下三个评价指标来测试各模型性能。

a) 均方根误差(RMSE)

$$RMSE = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(\hat{y}_i - y_i\right)^2} \quad (12)$$

b) 平均绝对误差(MAE)

$$MAE = \frac{1}{M}\sum_{i=1}^{M}\left|\hat{y}_i - y_i\right| \quad (13)$$

c) 平均绝对百分比误差(MAPE)

$$\mathrm{MAPE}=\frac{1}{M}\sum_{i=1}^{M}\left|\frac{\hat{y}_i-y_i}{y_i}\right|\times 100\% \tag{14}$$

其中：M 表示样本数量，y_i 表示第 i 个样本的实际交通速度，$\hat{y}_i$ 表示第 i 个样本的模型预测值。

3.3　实验结果分析

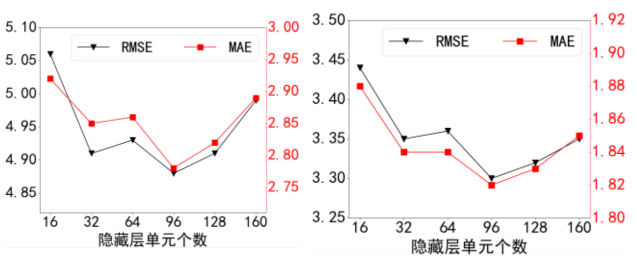
表 2 显示了 Ada-GGNN 和其他基线模型在两个数据集上进行 15、30 和 60 分钟交通预测的比较结果，从表 2 中可以得到以下观察结果：

基于神经网络的方法在两个数据集上的所有评估指标比传统的时间序列分析方法和机器学习方法(如 ARIMA 和 SVR)具有更低的预测误差。这些结果的原因是 ARIMA 和 SVR 模型的非线性建模能力有限。特别是 ARIMA 对数据的稳定性要求很高，但交通数据是复杂的，且影响因素较多，所以 ARIMA 呈现出最不理想的预测结果。相比之下，这些神经网络模型不仅擅长对非线性数据进行建模，还引入了拓扑结构获得交通路网的空间特征。因此，基于神经网络的方法表现出更好的性能。

Ada-GGNN 在两个数据集的长期和短期预测上实现了比其他基线模型更先进的预测性能。例如，对于 Los-loop 数据集上的 15 分钟交通预测任务，Ada-GGNN 在 RMSE 上比 GWN 提高了 2.21%，比 LSGCN 提高了 8.10%。类似的结果也出现在其他时间跨度预测任务和评估指标的比较中。这主要是 Ada-GGNN 模型的架构因素促进了这些改进。首先，Ada-GGNN 是一次输出多个时间步的预测结果，而 LSGCN 需要依靠先前的预测再生成下一时间步的预测，这样会导致预测误差的累积，所以在长期的预测结果上 Ada-GGNN 表现更优异。此外，Ada-GGNN 采用可学习的方式挖掘了交通路网中节点序列之间的关联关系，使得节点的嵌入表达更丰富，为模型引入了新的有用信息，有助于提高预测精度。最后，所提出的模型应用了循环聚合方法以实时融合方式学习时空嵌入，而不是单独的提取和延迟融合范式，这有助于学习细粒度的动态时空关系。

3.3.1 隐藏层单元数目的影响

通常，隐藏层单元的数目是导致过拟合的直接原因，因此隐藏层单元数目的设置在模型调优过程中至关重要。本实验从[16,32,64,96,128,160]中选择隐藏单元数量，并分析不同隐藏单元数量对模型预测精度的影响。如图 4 所示，横轴表示隐藏单元的数量，纵轴表示不同的度量指标。从图中可以看出，随着隐藏单元个数的增加，RMSE 和 MAE 指标先降低再增加，且当隐藏层单元个数为 96 时，Los-loop 和 PeMSD7 数据集的预测误差最低。这主要是因为隐藏单元个数过大时，模型复杂度较大，容易发生过拟合。



(a)Los-loop 数据集　　(b)PeMSD7 数据集

图 4　在两个数据集上不同隐藏层单元个数的影响

Fig. 4　Influence of different hidden units on both datasets

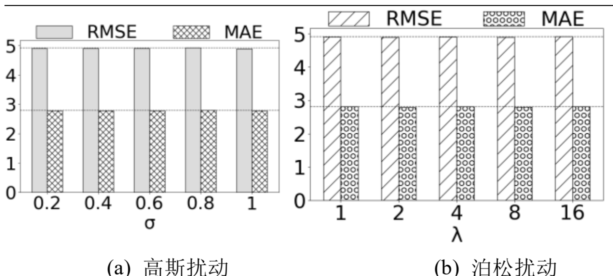
3.3.2 模型鲁棒性验证

在交通数据采集过程中，传感器难免会为数据引入噪声，因此，本实验通过扰动分析实验来测试 Ada-GGNN 的模型鲁棒性。如图 5 所示，在实验中添加了两种随机噪声，分别是高斯噪声和泊松噪声。高斯噪声服从高斯分布 $N\in(\mu=0,\sigma^2)$，其中方差 $\sigma\in[0.2,0.4,0.6,0.8,1]$；泊松噪声服从泊松分布 $P(\lambda)$，其中参数 $\lambda\in[1,2,4,8,16]$。从图中结果可以看出，Ada-GGNN 无论在何种噪声分布中，其评估指标的变化都很微小。因此，Ada-GGNN 可以适应噪声数据且具有很强的鲁棒性。

表 2　Ada-GGNN 与其他基准模型在 PeMSD7 和 Los-loop 数据集上的性能比较

Tab. 2　Performance comparison of Ada-GGNN and other baselines on pemsd7 and Los-loop datasets

Datasets	Models	15min			30min			60min		
		RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
PeMSD7	ARIMA	12.29	10.12	25.29%	12.30	10.12	25.30%	12.35	10.23	25.42%
	SVR	4.89	4.08	8.28%	5.90	4.66	9.77%	7.43	5.64	12.30%
	ASTGCN	3.41	1.94	4.50%	4.56	2.58	6.11%	6.11	3.31	8.39%
	T-GCN	3.45	1.93	4.41%	4.50	2.52	5.93%	6.08	3.23	8.32%
	A3T-GCN	3.40	1.89	4.30%	4.38	2.35	5.68%	5.82	3.39	8.19%
	LSGCN	3.37	1.92	4.33%	4.43	2.33	5.66%	6.17	3.48	8.38%
	GWN	3.39	1.89	4.32%	4.45	2.39	5.62%	5.96	3.15	7.86%
	Ada-GGNN	**3.30**	**1.82**	**4.21%**	**4.34**	**2.32**	**5.59%**	**5.46**	**2.91**	**7.18%**
Los-loop	ARIMA	10.05	7.71	20.84%	10.06	7.71	20.86%	10.07	7.72	20.90%
	SVR	5.65	3.86	9.33%	6.73	4.65	11.51%	8.03	4.22	14.65%
	ASTGCN	5.25	3.27	8.05%	6.14	3.76	9.68%	7.47	4.43	12.29%
	T-GCN	5.21	3.27	8.34%	6.12	3.81	9.96%	7.27	4.42	12.11%
	A3T-GCN	5.12	3.17	8.08%	6.10	3.73	9.98%	7.23	4.56	12.21%
	LSGCN	5.31	3.35	8.33%	6.24	3.82	9.95%	7.63	4.88	12.94%
	GWN	4.99	2.86	7.32%	5.93	3.26	8.81%	7.15	3.98	11.48%
	Ada-GGNN	**4.88**	**2.79**	**7.16%**	**5.83**	**3.18**	**8.68%**	**7.07**	**3.90**	**10.96%**



(a) 高斯扰动　　(b) 泊松扰动

图 5　高斯扰动和泊松扰动对 Ada-GGNN 在 Los-loop 数据集上的影响

Fig. 5　The influence of Gaussian perturbation and Poisson perturbation on the Ada-GGNN on the Los-loop dataset

3.3.3 自适应相关性矩阵的验证

为了验证空间传递模块中自适应相关性矩阵学习的有效性，设计了如下对比实验。其中，No_Ada 模型在 Ada-GGNN 的基础上去除了自适应时序相关性学习操作，仅采用预定义邻接矩阵来捕获交通路网的空间依赖性。在两个数据集上的 15 分钟和 60 分钟预测任务上的实验结果如表 3 所示。

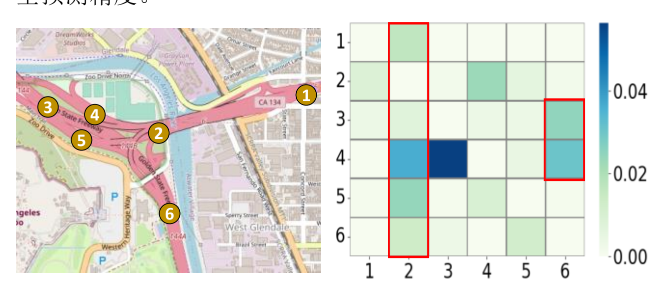
表 3　空间传递模块中时序相关性学习的影响

Tab. 3　Influence of learning time-series correlation in the SP module

Datasets	Models	15min			60min		
		RMSE	MAE	MAPE	RMSE	MAE	MAPE
PeMSD7	No_Ada	3.59	1.99	4.62%	5.99	3.18	7.98%
	Ada-GGNN	**3.30**	**1.82**	**4.21%**	**5.46**	**2.91**	**7.18%**
Los-loop	No_Ada	5.05	2.86	7.28%	7.59	4.08	11.55%
	Ada-GGNN	**4.88**	**2.79**	**7.16%**	**7.07**	**3.90**	**10.96%**

从表 3 的实验结果可以发现，在两个数据集的不同任务中，所提出的 Ada-GGNN 模型的预测性能明显优于 No_Ada 模型。这主要是因为 Ada-GGNN 模型不仅捕获了道路网络的空间关系，还学习到了各节点间的时序相关性，弥补了以往基于图神经网络的交通流预测方法仅含有单一的空间关系的缺陷。

同时，为进一步解释学习到的自适应相关性矩阵的作用，还进行了以下实验，即通过 Los-loop 数据集上部分节点的地理位置可视化和学习到的节点序列间相关性矩阵的热力图做进一步探索。图 6(a)是 6 个节点的地理位置图，图 6(b)是经过归一化处理后对应的 6 个节点序列之间的相关性矩阵热力图。首先，从图 6(b)可以明显看到，2 号节点所在列比其他列的高值点更多，这表明 2 号节点在预测时对图中其他节点更具影响，且该节点与其他节点序列相关性较高。其次，在预定义的固定邻接矩阵中，6 号节点由于距离 3、4 号节点太远，并不存在连接关系。然而从图 6(b)中能够发现，6 号节点对 3、4 号节点有较大影响且存在关联关系。这表明 Ada-GGNN 能够通过自适应相关性矩阵发现因距离原因在预定义邻接矩阵中丢失的节点连接，并为模型引入新的有用信息，提高模型预测精度。



(a)部分节点的地理位置可视化　(b)节点的自适应相关性矩阵热力图

图 6　学习到的自适应相关性矩阵

Fig. 6　The learned self-adaptive correlation matrix

3.3.4 模型计算花销

表 4 展示了 Ada-GGNN 和其他基于神经网络的基线模型在 Los-loop 数据集上关于计算时间花销的比较。在训练阶段，Ada-GGNN 比 ASTGCN 快 1.07 秒，比 GWN 快 8.65 秒，但是比 T-GCN、A3T-GCN 和 LSGCN 大概慢 2~3 秒左右，这主要是因为 T-GCN 仅使用一层图卷积和一个 GRU 进行流量预测，而 GWN 却具有复杂的模型结构，它在每个时空层都有残差连接，计算量更大。在测试集上，LSGCN 比其他模型花费了更多时间，这是因为 LSGCN 必须根据先前的预测来生成之后的预测结果，而其他模型都是一次生成多个时间步预测结果。因此考虑模型的预测精度和时间花销，所提出的 Ada-GGNN 仍然具有很强的综合表现。

表 4　Los-loop 数据集上的计算时间

Tab. 4　Computation time on the Los-loop dataset

Models	Computation Time	
	Training (s/epoch)	Testing (s)
ASTGCN	5.59	0.44
T-GCN	0.95	0.11
A3T-GCN	1.05	0.11
LSGCN	1.95	1.83
GWN	13.17	0.60
Ada-GGNN	4.52	0.41

4　结束语

本文针对交通流预测问题提出了一种新的端到端的自适应门控图神经网络(Ada-GGNN)模型。该模型利用可学习的自适应矩阵从数据中获取节点间的序列相关性，为模型引入了新的有用信息，并基于门控图神经网络来处理长期的预测问题。在两个公共交通数据集上，Ada-GGNN 取得了最优异的结果。考虑到实际的交通预测问题受到许多外部因素的影响，例如天气、温度、社会事件等，因此，在未来的工作中，将引入更多的额外因素作为特征，并探索在大图上应用 Ada-GGNN 的可扩展方法。

参考文献：

[1] 刘静，关伟. 交通流预测方法综述 [J]. 公路交通科技, 2004, 21 (3): 82-85. (Liu Jing, Guan Wei. A summary of traffic flow forecasting methods [J]. Journal of Highway and Transportation Research and Development, 2004, 21 (3): 82-85.)

[2] Williams B M, Hoel L A. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results [J]. Journal of Transportation Engineering, 2003, 129 (6): 664-672.

[3] Zivot E, Wang Jiahui. Vector autoregressive models for multivariate time series [M]// Modeling Financial Time Series with S-Plus®. New York, NY, USA: Springer, 2003: 369–413.

[4] 张晓利，贺国光，陆化普. 基于 K-邻域非参数回归短时交通流预测方法 [J] 系统工程学报, 2009, 24 (2): 178-183. (Zhang Xiaoli, He Guoguang, Lu Huapu. Short-term traffic flow forecasting based on K-nearest neighbors non-parametric regression [J]. Journal of Systems Engineering, 2009, 24 (2): 178 - 183.)

[5] Yu R, Li Yaguang, Shahabi C, *et al.* Deep learning: a generic approach for extreme condition traffic forecasting [C]// Proc of SIAM International Conference on Data Mining. 2017: 777–785.

[6] Shi Xingjian, Chen Zhourong, Wang Hao, *et al.* Convolutional LSTM network: a machine learning approach for precipitation nowcasting [C]// Proc of the 28th International Conference on Neural Information Processing Systems. Cambridge, MA: MIT Press, 2015: 802–810.

[7] Zhang Junbo, Zheng Yu, Qi Dekang. Deep spatio-temporal residual networks for citywide crowd flows prediction [C]// Proc of the 31st AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2017: 1655–1661.

[8] Li Yaguang, Yu R, Shahabi C, *et al.* Diffusion convolutional recurrent neural network: data-driven traffic forecasting [C/OL]/ Proc of the 6th International Conference on Learning Representations. [S. l.] : OpenReview. net, 2018. (2018) [2022-01-12]. https://openreview. net/forum?id=SJiHXGWAZ.

[9] Yu Bing, Yin Haoteng, Zhu Zhanxing. Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting [C]// Proc of the 27th International Joint Conference on Artificial Intelligence. California: IJCAI, 2018: 3634–3640.

[10] Guo Shengnan, Lin Youfang, Feng Ning, *et al.* Attention based spatial-temporal graph convolutional networks for traffic flow forecasting [C]// Proc of the 33rd AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2019: 922-929.

[11] Zhao Ling, Song Yujiao, Zhang Chao, *et al.* T-GCN: a temporal graph convolutional network for traffic prediction [J]. IEEE Trans on Intelligent Transportation Systems, 2020, 21 (9): 3848–3858.

[12] Cho K, Van Merrienboer B, Gulcehre C, *et al.* Learning phrase representations using RNN encoder-decoder for statistical machine translation [C]// Proc of EMNLP. Stroudsburg, PA: ACL, 2014: 1724-1734.

[13] Gori M, Monfardini G, Scarselli F. A new model for learning in graph domains [C]// Proc of IEEE IJCNN. Piscataway, NJ: IEEE Press, 2005: 729–734.

[14] Defferrard M, Bresson X, Vandergheynst P. Convolutional neural networks on graphs with fast localized spectral filtering [C]// Proc of the 30th International Conference on Neural Information Processing Systems. Red Hook, NY: Curran Associates Inc, 2016: 3837–3845.

[15] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks [C/OL]/ Proc of the 5th International Conference on Learning Representations. [S. l.] : OpenReview. net, 2017. [2022-01-12]. https://openreview. net/forum?id=SJU4ayYgl.

[16] Hamilton W L, Ying R, Leskovec J. Inductive representation learning on

large graphs [C]// Proc of the 31st International Conference on Neural Information Processing Systems. Red Hook, NY: Curran Associates Inc, 2017: 1025–1035.

[17] Velickovic P, Cucurull G, Casanova A, *et al.* Graph attention networks [C/OL]/ Proc of the 6th International Conference on Learning Representations. [S. l.] : OpenReview. net, 2018. (2018) [2022-01-12]. https://openreview. net/forum?id=rJXMpikCZ.

[18] Li Yujia, Zemel R, Brockschmidt M, *et al.* Gated graph sequence neural networks [C/OL]/ Proc of the 4th International Conference on Learning Representations. 2016. (2016) [2022-01-12]. http://arxiv. org/abs/1511. 05493.

[19] Zhang Yufeng, Yu Xueli, Cui Zeyu, *et al.* Every document owns its structure: inductive text classification via graph neural networks [C]// Proc of the 58th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA: ACL, 2020: 334–339.

[20] Yang Yiying, Wei Zhongyu, Chen Qin, *et al.* Using external knowledge for financial event prediction based on graph neural networks [C]// Proc of the 28th ACM International Conference on Information and Knowledge Management. New York, NY, USA: ACM Press, 2019: 2161–2164.

[21] Cui Zeyu, Li Zekun, Wu Shu, *et al.* Dressing as a whole: outfit compatibility learning based on node-wise graph neural networks [C]// Proc of WWW Conference. New York, NY, USA: ACM Press, 2019: 307–317.

[22] Wu Shu, Tang Yuyuan, Zhu Yanqiao, *et al.* Session-based recommendation with graph neural networks [C]// Proc of the 33rd AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2019: 346–353.

[23] Marino K, Salakhutdinov R, Gupta A. The more you know: using knowledge graphs for image classification [C]// Proc of IEEE CVPR. Los Alamitos, CA: IEEE Computer Society, 2017: 20–28.

[24] Li Ruiyu, Tapaswi M, Liao Renjie, *et al.* Situation recognition with graph neural networks [C]// Proc of IEEE ICCV. Los Alamitos, CA, USA: IEEE Computer Society, 2017: 4183–4192.

[25] Kingma D P, Ba J. Adam: a method for stochastic optimization [C/OL]/ Proc of the 3rd International Conference on Learning Representations. 2015. (2015) [2022-01-12]. http://arxiv. org/abs/1412. 6980.

[26] Smola A J, Schölkopf B. A tutorial on support vector regression [J]. Statistics and Computing, 2004, 14 (3): 199–222.

[27] Bai Jiandong, Zhu Jiawei, Song Yujiao, *et al.* A3T-GCN: attention temporal graph convolutional network for traffic forecasting [J]. ISPRS International Journal of Geo-Information, 2021, 10 (7): 485.

[28] Huang Rongzhou, Huang Chuyin, Liu Yubao, *et al.* LSGCN: long short-term traffic prediction with graph convolutional networks [C]// Proc of the 29th International Joint Conference on Artificial Intelligence. California: IJCAI, 2020: 2355-2361.

[29] Dauphin Y N, Fan A, Auli M, *et al.* Language modeling with gated convolutional networks [C]// Proc of the 34th International Conference on Machine Learning. [S. l.] : PMLR, 2017: 933-941.

[30] Wu Zonghan, Pan Shirui, Long Guodong, *et al.* Graph WaveNet for deep spatial-temporal graph modeling [C]// Proc of the 28th International Joint Conference on Artificial Intelligence. California: IJCAI, 2019: 1907-1913.